AFRL-RH-WP-TR-2012-0172

COMBATING VIGILANCE DECREMENTS IN A SUSTAINED ATTENTION TASK:

EXAMINATION OF TWO COGNITIVE INTERVENTION SCHEDULES FOR A SECONDARY TASK

Thomas R. Carretta
Guy A. French

Warfighter Interface Division

MARCH 2012

Distribution A: Approved for public release; distribution unlimited.

AIR FORCE RESEARCH LABORATORY
711 HUMAN PERFORMANCE WING,
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2012-0172 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

ROWE.ALLEN.J.1259846044

Digitally signed by ROWE.ALLEN.J.1259846044
DN: c=US, o=U.S. Government, ou=DoD, ou=PKI, ou=USAF, cn=ROWE.ALLEN.J.1259846044
Date: 2013.03.26 14:44:13 -04'00'

Allen J. Rowe
Work Unit Manager
Supervisory Control and Cognition

BARBATO.GREGORY.J.1230165846

Digitally signed by BARBATO.GREGORY.J.1230165846
DN: c=US, o=U.S. Government, ou=DoD, ou=PKI, ou=USAF, cn=BARBATO.GREGORY.J.1230165846
Date: 2013.03.29 12:25:02 -04'00'

Gregory J. Barbato
Chief, Supervisory Control and Cognition
Branch Warfighter Interface Division

FREEMAN.GEORGE.R.1019635976

Digitally signed by FREEMAN.GEORGE.R.1019635976
DN: c=US, o=U.S. Government, ou=DoD, ou=PKI, ou=USAF, cn=FREEMAN.GEORGE.R.1019635976
Date: 2013.04.18 09:54:30 -04'00'

G. Richard Freeman
Chief, Warfighter Interface Division
Human Effectiveness Directorate
711 Human Performance Wing

REPORT DOCUMENTATION PAGE	*Form Approved* OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
20-03-2012	Interim	15 May 2011 – 15 March 2012

4. TITLE AND SUBTITLE	
Combating Vigilance Decrements In A Sustained Attention Task: Examination Of Two Cognitive Intervention Schedules For A Secondary Task	5a. CONTRACT NUMBER NA
	5b. GRANT NUMBER NA
	5c. PROGRAM ELEMENT NUMBER 61102F
6. AUTHOR(S Thomas R. Carretta & Guy A. French	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER 71840919

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Air Force Materiel Command Air Force Research Laboratory 711 Human Performance Wing Human Effectiveness Directorate Warfighter Interface Division Supervisory Control and Cognition Branch Wright-Patterson AFB OH 45433	AFRL 711 HPW/RHCI 11. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-RH-WP-TR-2012-0172

12. DISTRIBUTION AVAILABILITY STATEMENT
Distribution A: Approved for public release; Distribution unlimited.

13. SUPPLEMENTARY NOTES
88 ABW Cleared 02/19/2013; 88ABW-2013-0794. Report contains color.

14. ABSTRACT
Results from previous studies (St. John & Risser, 2007, 2009) indicate the addition of a simple cognitive secondary task may mitigate vigilance decrements for a sustained attention task involving target acquisition. The effectiveness of the cognitive task increased when its onset was triggered by physiological indicators of inattention. The current study examined the generalizability of this methodology with a few modifications. A no intervention condition was added to provide a baseline and a short perceptual vigilance task (PVT) was added to examine the construct validity of the experimental task (ET). Finally, instead of using physiological indicators to trigger the intervention, a schedule was used that resembled that of the physiological intervention. Although vigilance decrements were observed for both the PVT and ET, only a weak relationship was observed between the two tasks. ET performance was not affected by the cognitive intervention. The apparent poor construct validity of the ET and failure to replicate previous findings cast doubts on the robustness of the cognitive intervention for mitigating performance decrements on real-world tasks, especially when its onset is not linked with physiological indicators of inattention.

15. SUBJECT TERMS
Sustained attention, vigilance

16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Allen Rowe
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER (*Include area code)*
U	U	U	UNL	24	

THIS PAGE INTENTIONALLY LEFT BLANK.

CONTENTS

Section	Page
List of Figures …………………………………………	iii
Preface …………………………………………	v
1. Summary …………………………………………	1
2. Introduction …………………………………………	1
3. Method …………………………………………	4
3.1 Participants …………………………………………	4
3.2 Measures …………………………………………	4
3.3 Equipment …………………………………………	5
3.4 Procedures …………………………………………	6
3.5 Analyses …………………………………………	7
4. Results …………………………………………	7
4.1 Objective Performance …………………………………………	7
4.2 Subjective Measures …………………………………………	9
5. Discussion …………………………………………	13
6. References …………………………………………	14

List of Figures

Figure		Page
1	Snapshots of the long (target) and short-bed (non-target) trucks used in the experimental task …………………………	7
2	Hit percent by time interval on the abbreviated perceptual vigilance task …………………………	8
3	Hit percent by time interval for the experimental target recognition task …………………………	9
4	Hit percent for each experimental condition by time interval for the experimental target recognition task …………………………	10
5	Mean subjective workload ratings after completion of the abbreviated perceptual vigilance task (PVT) and the experimental task (ET) …………….	11

6 **Mean subjective workload ratings after completion of the experimental task (ET) by experimental condition …………………………………………... 11**

7 **Mean subjective stress state ratings for the initial (pre-experiment) baseline and after completion of the abbreviated perceptual vigilance task (PVT) and the experimental task (ET) ………………………………….. 12**

8 **Mean subjective stress state ratings after completion of the experimental task (ET) by experimental condition …………………………………………….. 13**

Preface

This report describes activities performed in support of the development of interface technology designed to improve supervisory control of multiple information sources (711 HPW/RHCI), Work Unit 71840919, Supervisory Control Information Management Research (SCIMR). The authors thank Mr. Tony Ayala and Ms. Sarah Spriggs for software development support and Mr. John Flach for patience (and vigilance!) during data collection.

1.0 SUMMARY

Results from previous studies (St. John & Risser, 2007, 2009) indicate the addition of a simple cognitive secondary task may mitigate vigilance decrements for a sustained attention task involving target acquisition. The effectiveness of the cognitive task increased when its onset was triggered by physiological indicators of inattention. The current study examined the generalizability of this methodology with a few modifications. A no intervention condition was added to provide a baseline and a short perceptual vigilance task (PVT) was added to examine the construct validity of the experimental task (ET). Finally, instead of using physiological indicators to trigger the intervention, a schedule was used that resembled that of the physiological intervention. Although vigilance decrements were observed for both the PVT and ET, only a weak relationship was observed between the two tasks. ET performance was not affected by the cognitive intervention. The apparent poor construct validity of the ET and failure to replicate previous findings cast doubts on the robustness of the cognitive intervention for mitigating performance decrements on real-world tasks, especially when its onset is not linked with physiological indicators of inattention.

2.0 INTRODUCTION

The capability to capture video and/or still imagery through fixed and movable sensors has increased dramatically in recent years. Along with the increased ability to capture imagery comes the need to devote more resources to review and analyze it in order to improve situational awareness and enable timely decisions and actions as required. Although advances have been made in technologies such as automatic target recognition (Ratches, 2011), change detection (Lu, Mausell, Brondizio, & Moran, 2004), detection of anomalous behaviors (Booth, Jones, & Redding,2004; Hollywood, Snyder, McKay, & Boon, 2004), and facial recognition (Sinha, Balas, Ostrovsky, & Russell, 2006) to assist humans in the review of imagery and detection of potential threats, this process is still very labor intensive. As a result, sustained attention and vigilance have become critical components of human performance. Operator fatigue plays a crucial role in the ability of image analysts to effectively monitor imagery over long periods.

Operator fatigue has been addressed by applying work shifts, changing out the crew mid mission at scheduled intervals (typically 8 hours). However, vigilance decrements begin to affect human performance long before physical fatigue begins to affect operators. This is especially true for sensor operators and image analysts performing long missions. These missions require personnel to monitor sensor video imagery for long periods for targets and suspicious activity. The low rate of occurrence of such activities over time leads to performance decrements in terms of reduced target acquisition rates and longer response times. By building a better understanding of the factors that contribute to these vigilance decrements, we can identify and evaluate technologies designed to address and mitigate factors that contribute to vigilance decrements and augment human performance.

Perceptual vigilance (Davies & Parasuraman, 1982; Molloy & Parasuraman, 1996)) has been studied for several years, resulting in the identification and characterization of performance decrements in sustained attention tasks. Theorists have attempted to explain vigilance decrements as a function of arousal/motivation (Vroom, 1964; Yerkes & Dodson, 1908), workload/multiple resource theory (Wickens, 2002), and other factors. However, few effective mitigation technologies have been implemented to address the issue (Schroeder, Touchstone, Stern, Stoliarov, & Thackray, 1994; St John & Risser, 2007, 2009).

The approach used by St. John and Risser (2007) to mitigate the vigilance decrement effect combined aspects of arousal and resource theories (Arousal-Resource model). They introduced an intervention in the form of a secondary task designed to draw upon resources separate from those required to perform a primary visual vigilance task. The interventions included an auditory alarm "ring tone" that required sensory perception only and two auditory cognitive tasks that required participants to mentally reorder strings of 3 or 4 spoken digits. St. John and Risser hypothesized that the cognitive digit task interventions would arouse participants, replenish depleted resources, and re-engage them in the vigilance task. They also hypothesized that the cognitive tasks would be more effective than the simple alarm because they were more demanding and engaging. Participants performed a 45 minute laboratory vigilance task twice, once in a control condition without any intervention and once with one of the three interventions. In the intervention conditions, participants received the intervention whenever they missed a target. This intervention method served as a proxy for a closed-loop system in which participants would receive interventions whenever low attention was detected by physiological measures, prior to an actual miss. All three interventions significantly reduced misses by approximately 30%. Participants who showed greater vigilance decrements in the baseline (no intervention) condition showed more improvement from all interventions. That is, more vulnerable participants benefited most. The cognitive interventions performed as well as, but no better than, the simple alarm. The cognitive tasks also interfered with target detection performance on occasions when the interventions were active. However, the alarm was rated as more frustrating and less appropriate than either of the cognitive interventions.

French, Carretta, and Flach (2011) examined the utility of St John and Risser's (2007) Arousal-Resource approach for mitigating the effects of vigilance decrements in a reconnaissance, surveillance, and target acquisition (RSTA) task where a single operator was required to monitor sensor feeds from two remote sources. Contrary to St. John and Risser (2007), the interventions had no significant effect on reduction of the vigilance decrement. French et al. speculated that methodological differences between the two studies may have contributed to differences in results. To begin, there was a major difference in the primary tasks in the two studies. In St. John and Risser, participants performed a sensor monitoring task that involved detection of a single critical signal repeatedly over 45 minutes. The event exposure duration was 400 ms and the event rate was 1 per 2 seconds with a critical signal (target) rate of 3 per minute. The critical signal was a truck icon that was slightly larger than the non-critical signal (110 pixels vs. 100 pixels). All signals occurred in one of 6 fixed screen locations.

The primary task in the French et al. (2011) study was a step closer to a common sensor operator task. Rather than a single critical signal that occurred at fixed locations, 11 suspicious behaviors were used as critical signals that could occur anywhere in two sensor streams (i.e., displays). Further, the non-critical signals were "daily life" activities in an urban scene continuously presented. Thus, the non-critical signals overlapped with the targets. Finally, the critical signal rate in the French et al. study was 1 per 2.5 minutes across the two sensor streams with a median exposure time of 11.0 seconds. Despite these differences in event rate, exposure time, number of displays monitored, and observer uncertainty about spatial location of the critical signal, a vigilance decrement was observed. The difference between the studies occurred in the effectiveness of the interventions. Contrary to St. John and Risser (2007), the intervention in the French et al. study did not significantly affect performance relative to the no intervention control condition.

One possible explanation for the difference in findings for the two studies is the intervention schedule. St. John and Risser (2007) implemented their interventions with simulated physiological monitoring where the interventions were triggered by a missed target. French et al. (2011) used a simpler approach, where the interventions were introduced on a quasi-random schedule. Such an approach, if effective, would eliminate the need for physiological monitoring to achieve performance benefits. Unfortunately, the intervention did not produce the desired effect in the French et al. study. Whether the reason for different results lies in the differences in the nature of the experimental tasks, the intervention schedule, or some interaction of the two, it appears that physiological monitoring may be necessary to achieve the performance benefits of a system based mitigation strategy.

St. John and Risser (2009) subsequently conducted a study that examined the effectiveness of their cognitive intervention task for mitigating the effects of vigilance decrements under two conditions. Contrary to St. John and Risser (2007) and French et al. (2011), there was no "no intervention" baseline condition. In the first condition, the cognitive intervention was implemented on a quasi-random schedule, occurring once every 2 minutes +/- 30 seconds. In the other condition, the cognitive intervention was triggered by physiological indicators of inattention (i.e., eye movement, head movement, and EEG). While participants in both conditions exhibited a vigilance decrement as indicated by target misses over the 45-minute session, those in the physiological condition had 17% fewer target misses overall compared to the random schedule condition. St. John and Risser (2009) noted that the improvement in performance was not due to the intervention alone, but the intervention schedule. Roughly the same number of interventions occurred for both conditions. However, whereas the interventions were more or less evenly spaced in the quasi-random schedule condition, they started out at a low level in the physiological condition and became more frequent as the task lengthened till they reached a steady state after about 20 minutes.

The current study examined the utility of a variable intervention schedule that mimics the physiological-based pattern used by John and Risser (2009) with the addition of a no intervention control group to provide a baseline. An artificial vigilance task (Temple, Warm, Dember, Jones, LaGrange, & Matthews, 2000) was added as a pre-test to determine participants' vigilance ability and serve as a control variable in the analyses. We felt it was necessary to determine the replicability of the effect obtained by St. John and Risser (2009) before applying their approach to a more realistic target detection task where a single operator is required to monitor sensor feeds from two remote sources.

3.0 METHOD

3.1 Participants

Participants were 27 adult volunteers (16 males and 11 females). Most were military or civilian employees stationed at Wright-Patterson, AFB, OH. They ranged in age from 19 to 61, with a mean of 29.8 years. All participants reported they were in good health, with normal visual acuity after correction and no problems with peripheral vision or color blindness. Most reported some prior simulator experience (63%) and video game experience (67%). Participants reported getting between 6 and 9 hours of sleep the night before the experiment with a mean of 7.1 hours. Scores on the Epworth Sleepiness Scale (ESS) ranged from 2 to 17 with a mean of 6.4.

3.2 Measures

Several measures were collected prior to, during, and following the experimental target detection task. These are described in the following paragraphs.

3.2.1 Objective task performance. Objective measures of performance for the Experimental Task (ET) included proportion of hits and correct rejections and number of false alarms.

3.2.2 Perceptual Vigilance Task (PVT). In this task (Temple, Warm, Dember, Jones, LaGrange, & Matthews, 2000), participants monitored the presentation of 8- by 6-mm light grey capital letters that consisted of 'O', 'D', and a backwards 'D' centered on a video display screen. The letters were presented in 24-point type using an AvantGarde font and were exposed for 40 milliseconds against a visual mask that consisted of unfilled circles on a white background (see Figure 1). Participants were instructed to use the mouse to indicate when the target letter 'O' was presented. Responses were scored as hits, misses, and false alarms. Mean response time for hits and false alarms also were computed.

3.2.3 Biographical questionnaire. This questionnaire was used to collect information in order to characterize the sample and assist in interpretation of participants' performance on the target detection task. Items elicited information about participants' sex, age, general health, wellbeing, previous experience with simulator-type environments, previous experience with video games, and whether they had vision correctable to 20/20 acuity and normal peripheral and color vision.

3.2.4 NASA Task Load Index (NASA TLX). The NASA TLX (Hart & Staveland, 1988) is a subjective workload assessment measure that allows users to evaluate their interactions with human-machine systems. A multidimensional weighting procedure can be used to derive an overall workload score based on weighted averages of ratings on 6 subscales: Mental Demand, Physical Demand, Temporal Demand, Effort, Performance, and Frustration. In the present study overall workload was computed as the average of the values for the 6 subscales.

3.2.5 Short Stress State Questionnaire (SSSQ). The SSSQ (Helton, 2004) is an abbreviated version of the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 1999). It consists of four subscales: mood state, motivation, thinking style, and thinking content. Scores for three factors are derived from the subscales: task engagement, distress and worry. Additionally, one component of the SSSQ assesses participants' perceptions of their physical and mental workload. One version of the SSSQ was administered before task performance and the other was administered after task performance.

3.2.6 Epworth Sleepiness Scale (ESS). The Epworth Sleepiness Scale (ESS) is an eight-item questionnaire that is used as a simple method for measuring daytime sleepiness in adults. Total ESS score can range from 0 to 24. ESS scores are an indicator of average sleep propensity (ASP) for the 8 situations making up the scale items and are not synonymous with fatigue or tiredness. Factor analysis indicates the scale measures only one factor and is reported to have high internal consistency with a Cronbach's $\alpha = 0.88$ (Johns, 1992).

3.3 Equipment

Participants viewed simulated imagery on a 24-inch monitor and used a mouse to indicate target detection decisions. A portion of the 24-inch monitor was blocked off to emulate the 17-inch display set at a resolution of 1024 by 768 pixels as used by St. John and Risser (2007, 2009).

3.4 Procedures

The study began with a pre-briefing regarding research objectives and procedures, informed consent, and collection of demographic data about factors that may be related to performance on the target detection task (e.g., age, gender, and experience with similar tasks).

Participants then completed the SSSQ to provide a baseline of their stress level. Prior to examination of the effect of intervention schedule, a baseline was established for participants' perceptual vigilance using a 12 minute task based on a method described by Temple et al. (2000). These perceptual vigilance task (PVT) data were used to examine the construct validity of the experimental task (ET). If high construct validity was established between the PVT and ET, the data could be used in the analyses of the ET performance as a covariate to control for differences in perceptual vigilance among the study participants. Following the 12-minute PVT, participants completed the SSSQ to assess changes from their baseline (pre-PVT) level. They also completed the NASA TLX to assess their subjective workloads. There was a 10-minute break following completion of the PVT and questionnaires.

Participants were randomly assigned to one of the three experimental conditions: a "No Intervention" control condition, a Quasi-Random Intervention condition, or an Emulated Physiological Intervention condition. As with the perceptual vigilance task, prior to and following the experimental task, participants completed the SSSQ and NASA TLX following the experimental task.

All participants completed several training trials that consisted of a 2 minute task demonstration with auditory feedback and a target rate higher than that used in the rest of the trials (1 in 3). The high target rate during the demonstration was intended to provide participants greater opportunity to observe the target while learning the task. Participants then completed one to six 3-minute practice sessions during which the target rate was the same as that used in the experimental task (1 in 10). Next, the intervention task was explained followed by completion of a 10-trial practice session with feedback on the intervention task alone.

Participants were instructed to press the Space bar when they detected a target. Feedback was provided on hits, misses, and false alarms during the practice trial. After a short break, the 45-minute experimental trial followed during which performance feedback was given only for the intervention task. Following the experimental trial, participants completed the post-trial SSSQ and NASA TLX to assess their subjective stress (mood state) and workload.

The participants' task was to monitor simulated snapshots from a Remotely Piloted Aircraft (RPA) flying along a highway and to designate targets as they were detected. The event exposure duration was 400 ms and the event rate was 1 per 2 seconds with a critical signal (target) rate of three per minute. The critical signal was a truck icon that was slightly larger than the non-critical signal. All signals occurred in one of 6 fixed screen locations. Participants were responsible for designating targets and responding to interventions, depending on the experimental condition. The experimental trial lasted about 45 minutes. Figures 1 illustrates the displayed imagery as viewed by study participants.

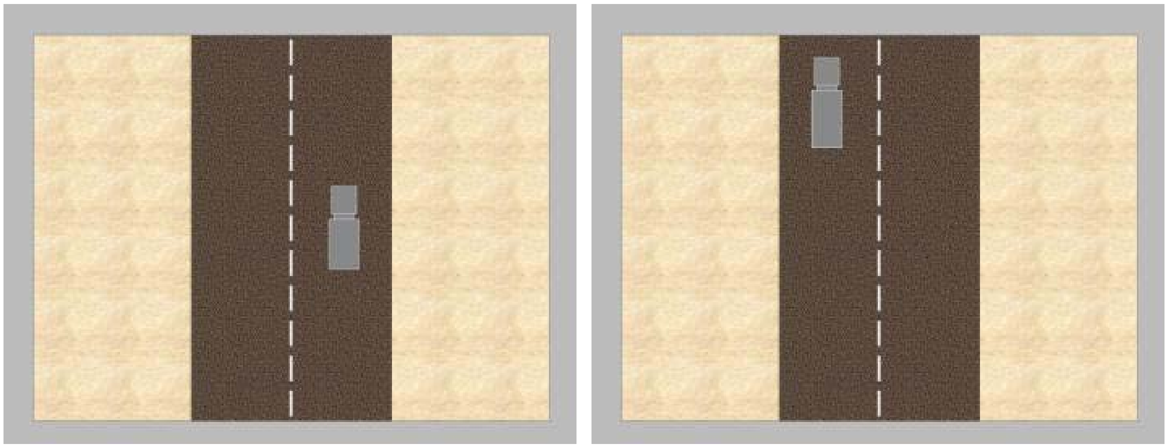


Figure 1. Snapshots of the short-bed (non-target) and long (target) trucks used in the experimental task.

3.5 Analyses

In order to examine trends over time, the abbreviated perceptual vigilance task (PVT) was divided into six 2-minute time intervals and objective performance (hits, correct rejections, false alarms) scores were calculated for each interval. Repeated measures ANOVAs were used to examine trends in performance over the six intervals. This technique also was used for the experimental task (ET), but nine 5-minute intervals were used. Correlational analyses examined the relations between the PVT and ET measures. ANOVAs were used to examine the effects of intervention condition (No Intervention Control, Random, Physiologically-Based) on ET performance measures.

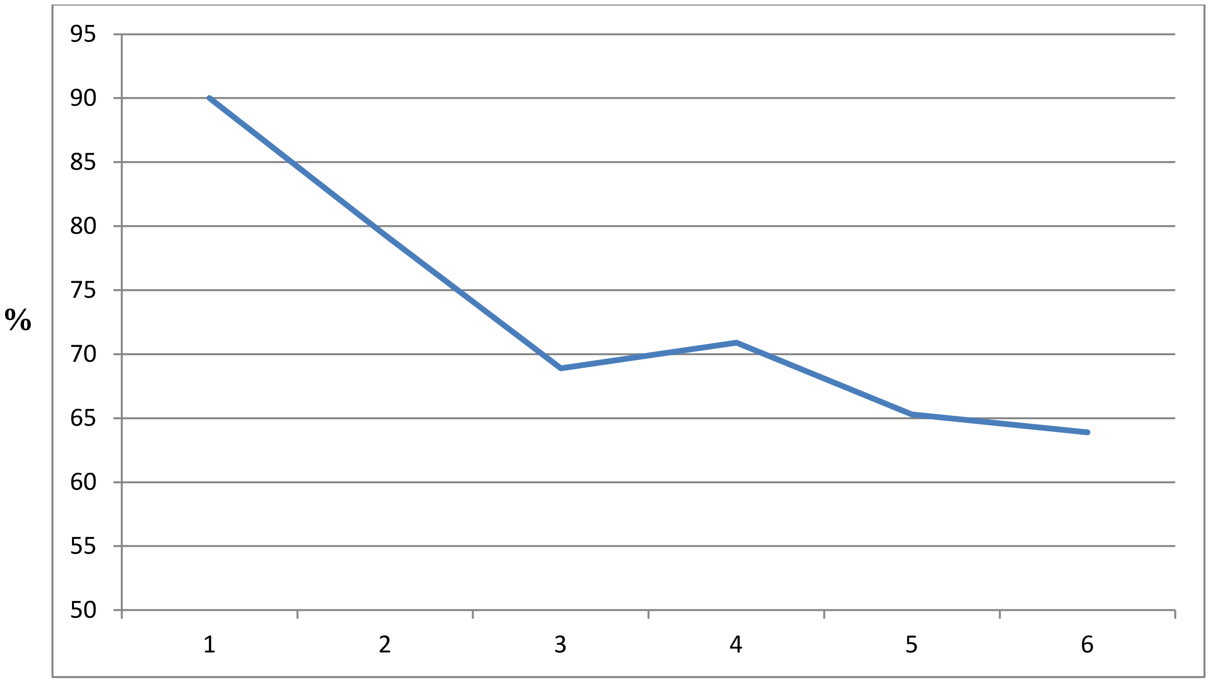
NASA TLX means were examined to compare subjective workload levels for the ET Control and Intervention conditions using t-tests. Pre-PVT SSSQ scores served as a baseline by which to evaluate changes in subjective mood state following the PVT and ET using related samples t-tests. All analyses used a .05 Type I error rate.

4.0 RESULTS

4.1 Objective Performance

As shown in Figure 2, the proportion of hits on the abbreviated PVT decreased over the six time intervals ($F(5, 130) = 11.97$, $p \leq .001$). There was no trend for proportion of correct rejections over time (% correct rejections = 98.0, 87.7, 97.5, 97.0. 96.8, and 96.7) ($F(5,130) = 1.58$, ns). There was a trend towards increased number of false alarms over the six time intervals (n false alarms = 1.9, 2.3, 2.4, 2.8. 3.0, 3.2), but it was not statistically significant ($F(5, 130) = 1.60$, ns).

Results for the ET were similar to those seen for the PVT. As shown in Figure 3, the proportion of hits on the ET generally decreased over time ($F(8, 192) = 8.13$, $p \leq .001$). Neither the proportion of correct rejections ($F(8, 192) = 1.24$, ns) nor the number of false alarms ($F(8, 192) = 1.42$, ns) were affected by time interval.



Time Interval

Figure 2. Hit percent by time interval on the abbreviated perceptual vigilance task. The 12-minute task was divided into six 2-minute intervals for analysis.

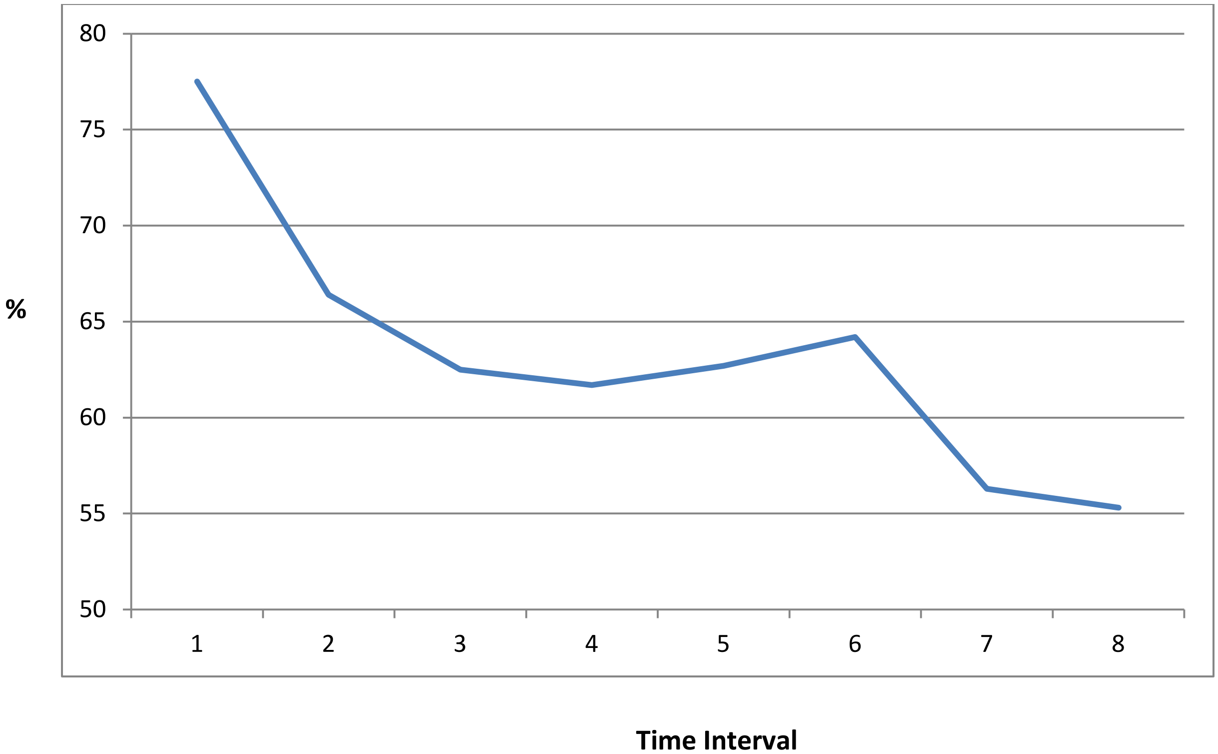
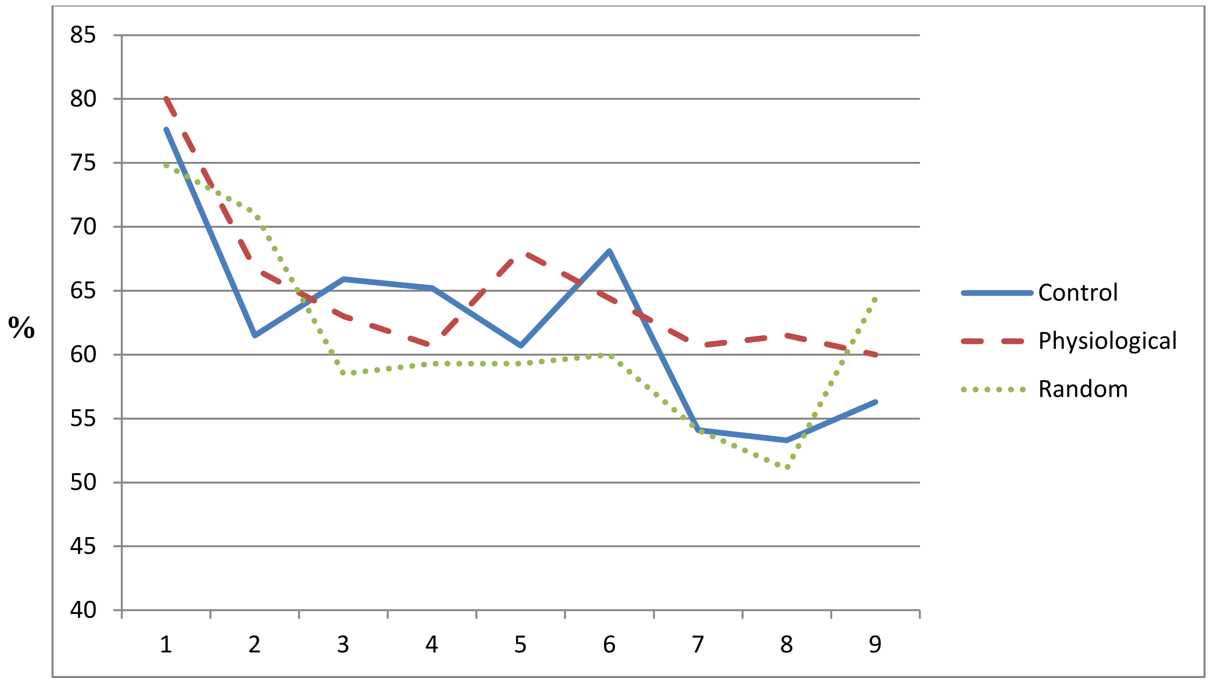


Figure 3. Hit percent by time interval for the experimental target recognition task. The 45-minute task was divided into nine 5-minute intervals for analysis.

Overall, performance was somewhat better on the abbreviated PVT than on the ET. The mean hit percents were 73.1% (PVT) and 63.0% (ET) and the correct rejection percents were 97.2% (PVT) and 96.7% (ET). Contrary to expectations, performance on the PVT and ET were not related. The correlations for like measures from the two tests were: hit percent (-.030), correct rejection percent (-.044), and number of false alarms (-.028). A factor analysis using oblique rotation was conducted to further examine the relations between the PVT and ET scores (hit percent, correct rejection percent, and number of false alarms). Two factors emerged that accounted for 78.3% of the variance of the scores, one factor each representing the PVT and ET. The correlation between the factors was 0.166.

As previously discussed, the proportion of hits on the ET significantly decreased over time ($F(8, 192) = 8.13$, $p \leq .001$). However, there was no significant effect for experimental condition (No Intervention Control, Random Intervention, Physiologically-based Intervention Schedule) ($F(2, 24) = 0.10$, ns). The mean hit percent for the three conditions were: Control (62.5%), Random Intervention (61.4%), and Physiologically-based Intervention (65.0%). See Figure 4.



Time Interval

Figure 4. Hit percent for each experimental condition by time interval for the experimental target recognition task.

4.2 Subjective Measures

Subjective workload was measured after completion of the abbreviated PVT and again after completion of the ET. As shown in Figure 5, most of the mean workload ratings for both tasks were in the low to moderate range. The highest rating was for Temporal Demand (TD) for the PVT with a mean of 71.1. The high TD score was not surprising as the pace of stimulus presentation for the PVT is high (57 stimuli per minute).

Paired-samples t-tests were conducted to determine whether workload ratings differed for the PVT and ET. Two-tailed tests were used as we had no preconceived hypotheses about the relative difficulty of the tasks. The only statistically significant effect occurred for the Temporal Demand (TD) scale ($t(26) = 2.83$, $p \leq .01$) where the PVT (mean = 71.1) was judged more difficult than the ET (mean = 52.4). There was a marginal difference for Mental Demand ($F(28) = -1.71$, $p \leq .0$) where the PVT (57.7) was judged to be less difficult than the ET (68.8).

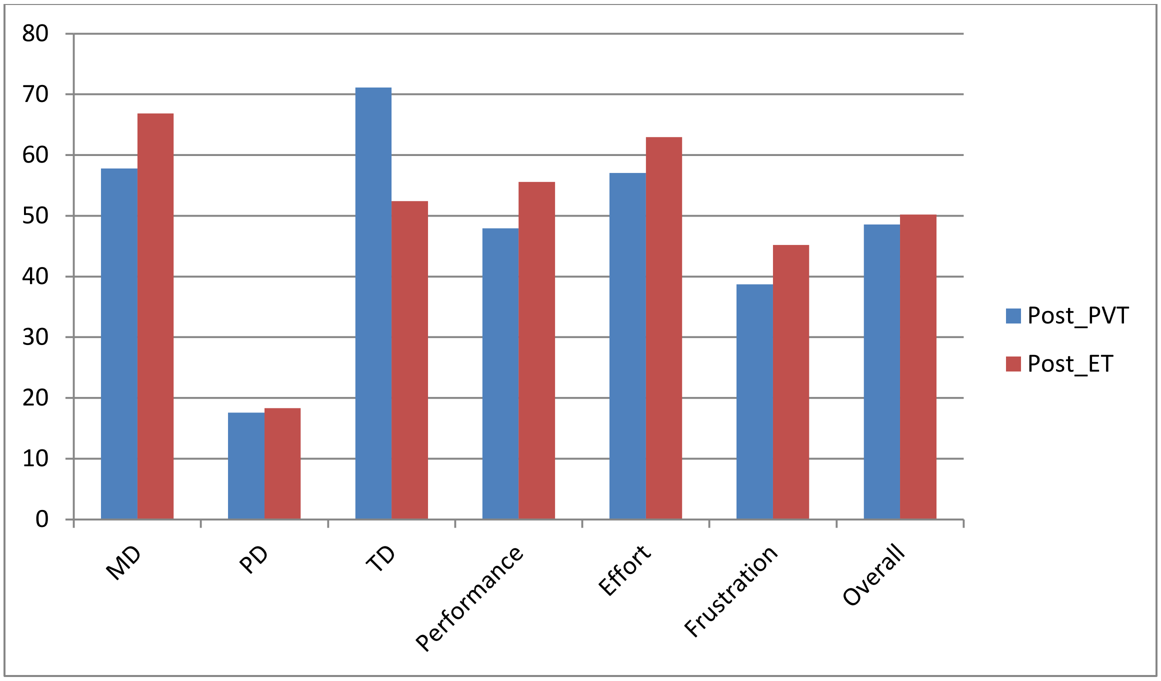


Figure 5. Mean subjective workload ratings after completion of the abbreviated perceptual vigilance task (PVT) and the experimental task (ET).

Figure 6 summarizes the workload ratings for the three experimental groups following completion of the ET. None of the mean differences were statistically significant.

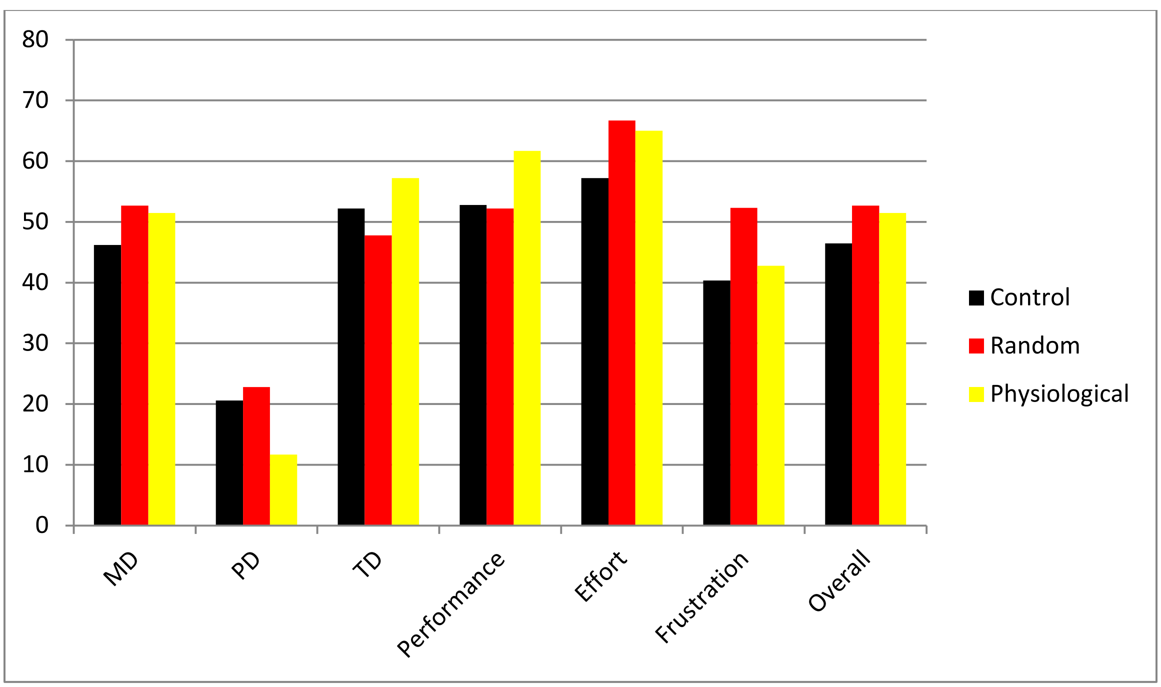


Figure 6. Mean subjective workload ratings after completion of the experimental task (ET) by experimental condition.

The Short Stress State Questionnaire (SSSQ) was administered prior to the PVT to provide an initial baseline to evaluate changes following the PVT and following the ET. Figure 7 summarizes scores on the SSSQ. It was expected that stress would increase relative to the baseline following the PVT and ET (i.e., lower Engagement, higher Distress and Worry). Statistically significant changes in stress state were observed in the expected direction for five of the six comparisons with the initial baseline. Engagement scores decreased following the PVT ($t(26) = 1.99$, $p \leq .05$) and ET ($t(26) = 6.83$, $p \leq .001$), while Distress increased following the PVT ($t(26) = -6.04$, $p \leq .001$) and ET ($t(26) = -4.76$, $p \leq .001$). Worry increased following the ET ($t(26) = -2.27$, $p \leq .025$), but contrary to expectations decreased following the PVT.

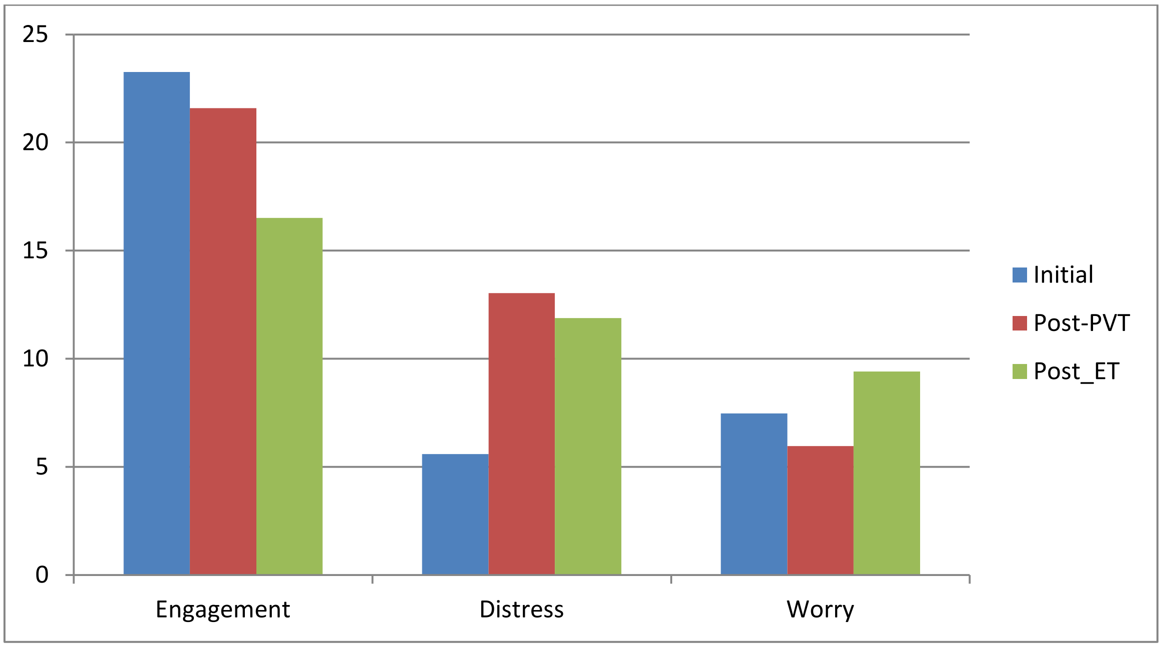


Figure 7. Mean subjective stress state ratings for the initial (pre-experiment) baseline and after completion of the abbreviated perceptual vigilance task (PVT) and the experimental task (ET).

Figure 8 summarizes the mean post-ET SSSQ scale scores for the three intervention conditions. Analyses of variance results indicated no statistically significant effects for any of the scales.

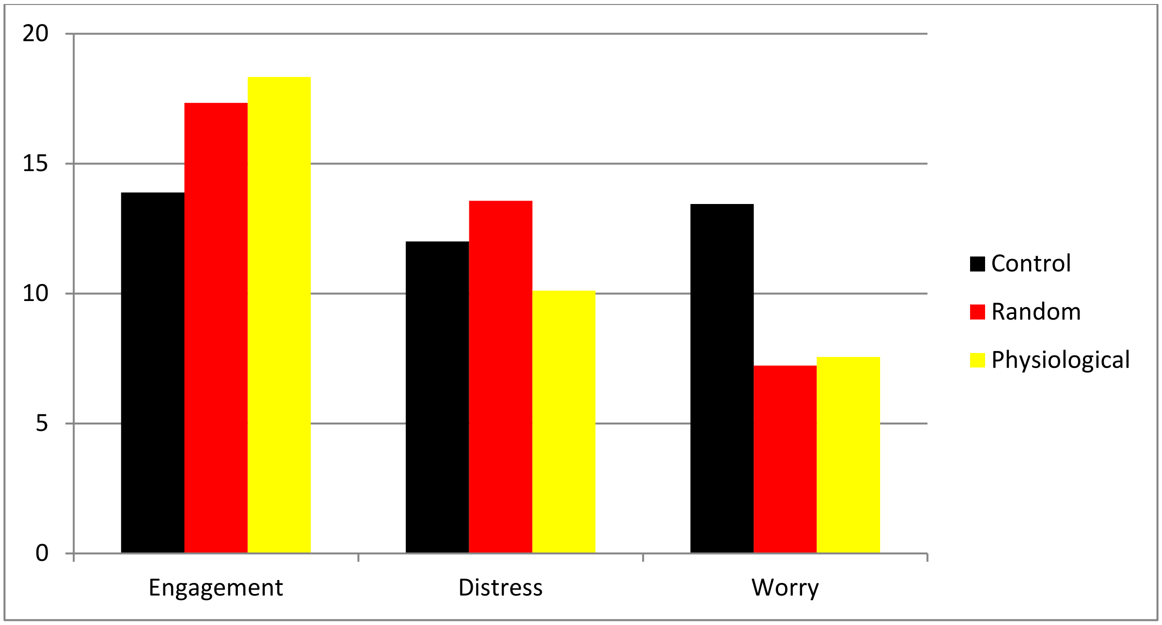


Figure 8. Mean subjective stress state ratings after completion of the experimental task (ET) by experimental condition.

5.0 DISCUSSION

The low correlations between similar measures of performance from the abbreviated PVT and ET are perplexing. Each measure was designed to assess sustained attention and vigilance. They use similar formats where a limited set of targets and distracters are used and displayed for short intervals. Further, the internal consistency reliabilities for each of the scores were good (PVT: hits. .90, correct rejections, .96, false alarms, .96; ET: hits, .94; correct rejections, .92, false alarms, .91). Despite similar appearance and task requirements and good reliability of the measures of performance, factors representing the two tasks exhibited only a weak relationship (r = 0.166).

Contrary to St. John and Risser (2007), inclusion of a cognitive intervention did not improve performance beyond that observed for a No Intervention Control. Further, contrary to St. John and Risser (2009) the difference in performance for the Random and Physiologically-Based Intervention schedules was not statistically significant. St. John and Risser reported hit percents of 64% and 70% for their Random and Physiologically-Based Intervention schedules compared with 61.4% and 65.0% in the current study. The relatively smaller improvement in the current study for the Physiologically-Based schedule may be because its implementation was not triggered by a missed target (St. John & Risser, 2007) or by indicators of fatigue/inattention (St. John & Risser, 2009). Even so, this does not explain why performance for the two intervention conditions was not better than that observed for the No Intervention Control (62.5%) condition in the current study.

It is unknown whether the difference in results for St. John and Risser (2007, 2009) and the French et al (2011) and the current study lie in the trigger mechanism for the cognitive intervention. Implementation of a vigilance decrement mitigation intervention in an operational setting would be greatly facilitated if it were not required to link it with physiological indicators of inattention in order to achieve effectiveness. Regardless, the low correlations between scores on the abbreviated PVT and the experimental task suggests that they are not measuring the same constructs (i.e., do not share construct validity). Further, the failure to replicate previous findings cast doubts on the robustness of the effectiveness of a simple cognitive intervention task for mitigating vigilance decrements in performance on real-world tasks that require sustained attention.

6.0 REFERENCES

Booth, D. M., Jones, K. L., & Redding, N. J. (2004). *Detection of anomalous behavior using video surveillance: Potential directions for the Analysts' Detection Support System,* DSTPO-TR-1611. Edinburgh, South Australia: DSTO Information Sciences Laboratory, Intelligence, Surveillance, and Reconnaissance Division.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

French, G. A., Carretta, T. R., & Flach, J. K. (2011). Vigilance decrements in a sustained attention task: Examination of a mitigation strategy. *Proceedings of the 16th International Symposium on Aviation* Psychology, Dayton, OH.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA TLX (task load index): Results of empirical and theoretical research in human mental workload. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). New York: Elsevier.

Helton, W. S. (2004). Validation of a short stress state questionnaire. *Proceedings of the 48th Meeting of the Human Factors and Ergonomics Society.*

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*(1), 18-25.

Hollywood, J., Snyder, D., McKay, K., & Boon, J. (2004). *Out of the ordinary: Finding hidden threats by analyzing unusual behavior.* Santa Monica, CA: RAND Corporation.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep, 15*, 376-381.

Lu, D., Mausell, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 23,* 2365-2407.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., &Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion, 2,* 315-340.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S. E., Huggins, J., & Falconer, S. (1999). Validation of a Comprehensive Stress State Questionnaire: Towards a State "Big Three"? In. I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe, 7,* Tilburg University Press: Tilburg, The Netherlands

Molloy R, & Parasuraman, R. (1996). Monitoring an automated system for a single failure: vigilance and task complexity effects. *Human Factors, 38*(2),311-322.

Ratches, J. A. (2004). Review of current aided/automatic target acquisition technology for military target acquisition tasks. *Optical Engineering, 50,* 072001. http://dx.doi.org/10.1117/1.3601879.

Schroeder, D. J., Touchstone, R. M., Stern, J, A., Stoliarov, N., & Thackray, R. (1994). *Maintaining vigilance on a simulated ATC monitoring task across repeated sessions,* DOT/FAA/AM-94/6. Washington, DC: U.S. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine.

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE Meeting, 64 (11),* 1948-1962.

St. John, M., & Risser, M. R. (2007). Mitigating the vigilance decrement: Perceptual versus cognitive interventions. *Proceedings of the 4th International Conference on Augmented Cognition,* Baltimore, MD.

St. John, M., & Risser, M. R. (2009). Sustaining vigilance by activating a secondary task when inattention is detected. *Proceedings of the 53rd Meeting of the Human Factors and Ergonomics Society,* Santa Monica, CA.

Temple, J. C., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors, 42*(2), 183-194.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*, 150-177.

Wilson, G. F., Russell, C., & Caldwell, J. (2006). *Performance and psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty*, AFRL-HE-WP-TP-2006-0095. Wright-Patterson AFB, OH: Human Effectiveness Directorate, Human Effectiveness Division.

Yerkes, E. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459-482.